SUPPLEMENTAL MATERIALS

A Genome-wide Polygenic Score, Clinical Risk Factors, and Long-term Trajectories of Coronary Artery Disease

George Hindy, MD, PhD,[1] Krishna G. Aragam, MD, MS,[2,3,4] Kenney Ng, PhD,[5] Mark Chaffin, MS,[2] Luca A. Lotta, MD, PhD,[6] Aris Baras, MD,[6] Regeneron Genetics Center, Isabel Drake, PhD,[7] Marju Orho-Melander, PhD,[7] Olle Melander, MD, PhD,[7] Sekar Kathiresan, MD,[2,4,8] Amit V. Khera, MD,[2,3,4,8]

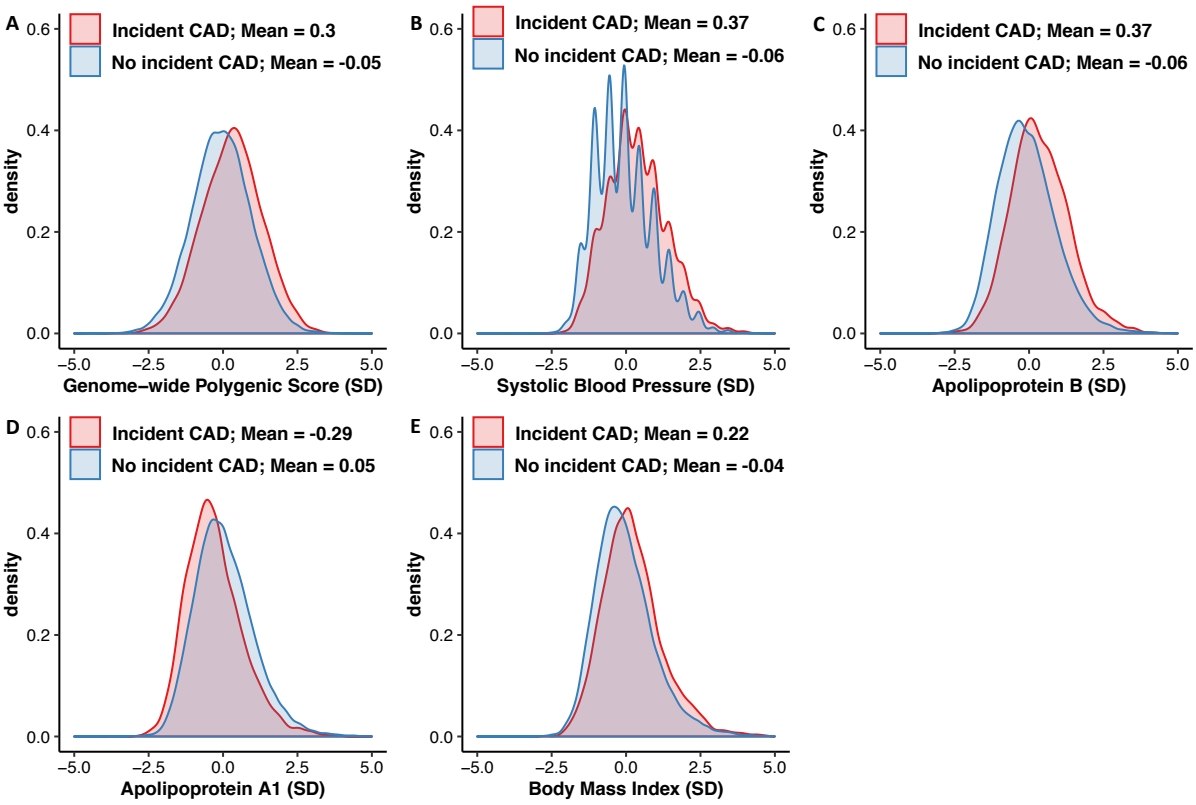
[1]Department of Population Medicine, Qatar University College of Medicine, QU Health, Doha, Qatar
[2]Program of Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, USA
[3]Center for Genomic Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA
[4]Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA
[5]Center for Computational Health, IBM Research, Cambridge, Massachusetts, USA
[6]Regeneron Genetics Center, Tarrytown, New York, USA
[7]Department of Clinical Sciences in Malmö, Lund University, Malmö, Sweden
[8]Division of Cardiology, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA

Running title: Polygenic and Clinical Risk Factors for Coronary Disease

Online Methods:

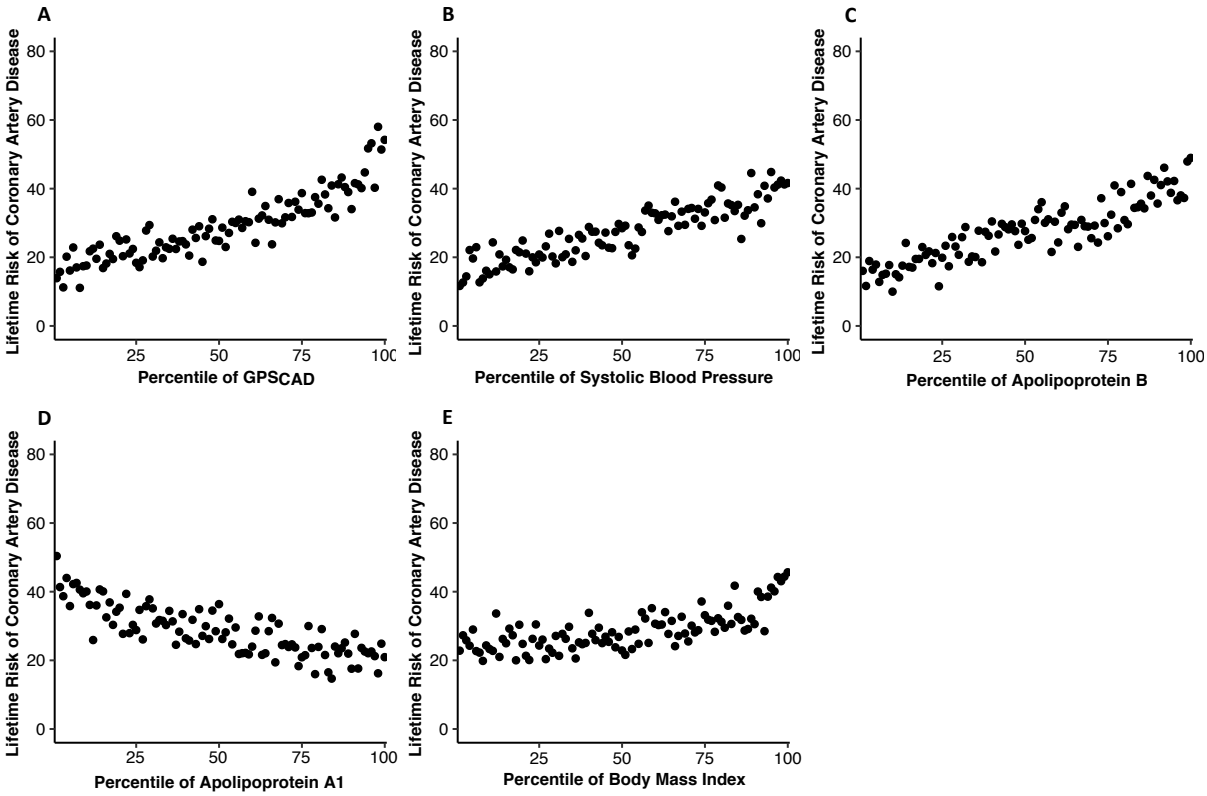
In the Malmö Diet and Cancer study, ancestry was inferred using a common set of genotyped single nucleotide polymorphisms (SNPs) with HapMap3. The set was filtered to 103,449 common, high quality SNPs with minor allele frequency above 5% and call rate above 99%. After merging both the Malmö Diet and Cancer study with samples from HapMap3, principal components were calculated and a kernel density estimator was trained for each ancestral superclass with the principal components of the HapMap3 samples. Then the likelihood of each sample being from the different ancestral superclasses based on the kernel density estimator was calculated.

Figure I. Distribution of the Genome-wide Polygenic Score and Clinical Risk Factors Stratified by Incident Coronary Artery Disease in the Malmö Diet and Cancer Study



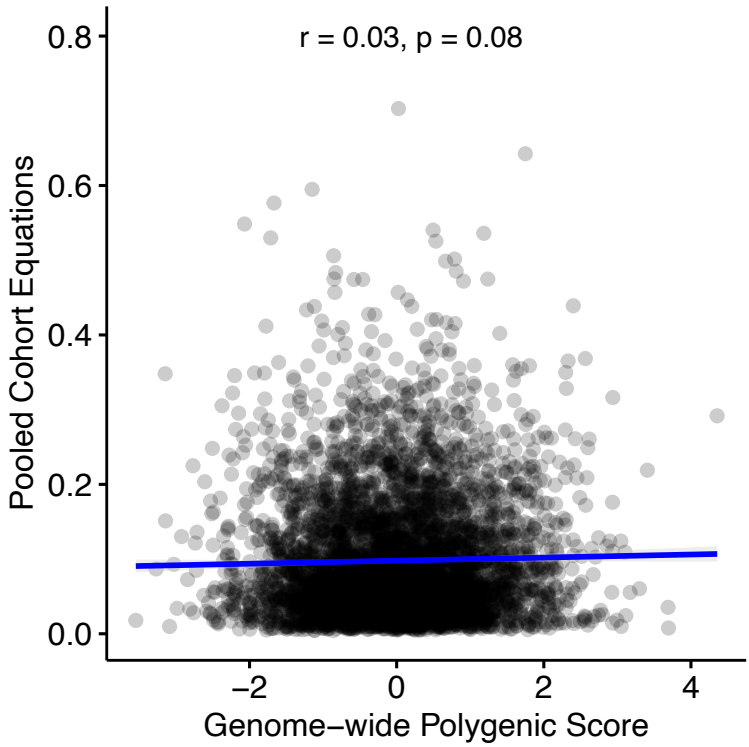
Density plots stratified according to incident coronary artery disease events are depicted for **A**) genome-wide polygenic score (GPS_{CAD}), **B**) systolic blood pressure, **C**) apolipoprotein B, **D**) apolipoprotein A1 and **E**) body mass index. For each risk factor, the x-axis is displayed according to z-score (scaled to mean of 0 and standard deviation of 1) to facilitate comparisons.

Figure II. Lifetime Risk of Coronary Artery Disease According to Percentile of the Genome-wide Polygenic Score and Clinical Risk Factors



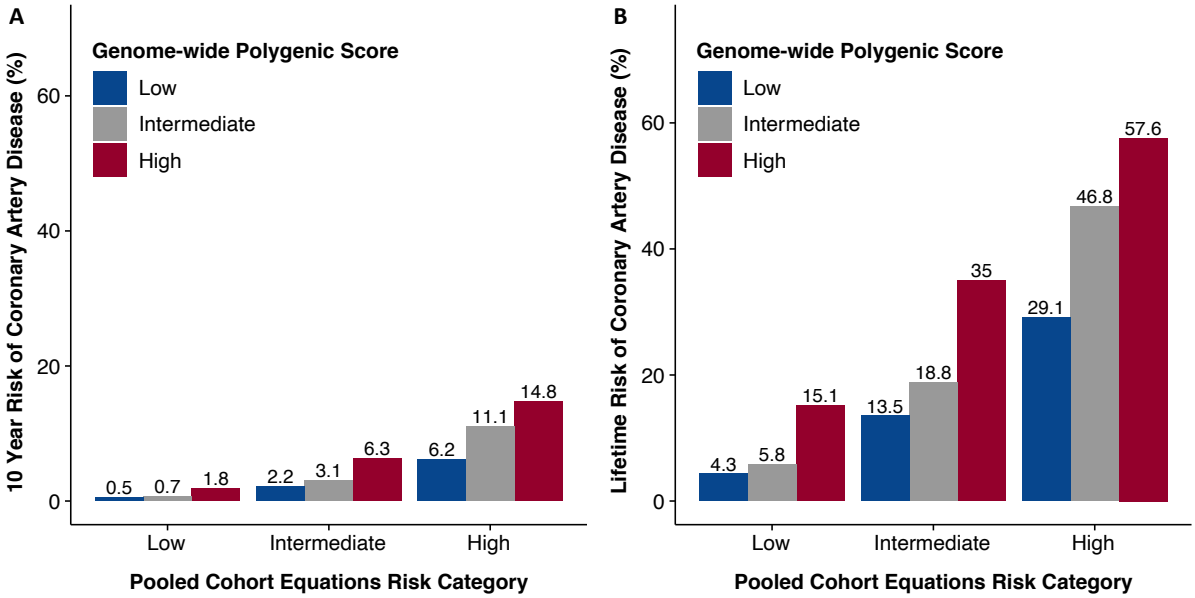
Lifetime risk of coronary artery disease in the Malmö Diet and Cancer Study − defined as cumulative risk by 90 years of age − was determined using Cox proportional hazard models with age as the underlying time scale and standardized for sex. Values are depicted for **A**) the genome-wide polygenic score for coronary artery disease (GPS_{CAD}), **B**) systolic blood pressure (SBP), **C**) apolipoprotein B (ApoB), **D**) apolipoprotein A1 (ApoA1) and **E**) body mass index (BMI).

Figure III. Correlation between the Genome-wide Polygenic Score and 10-year risk Predicted by the Pooled Cohort Equations



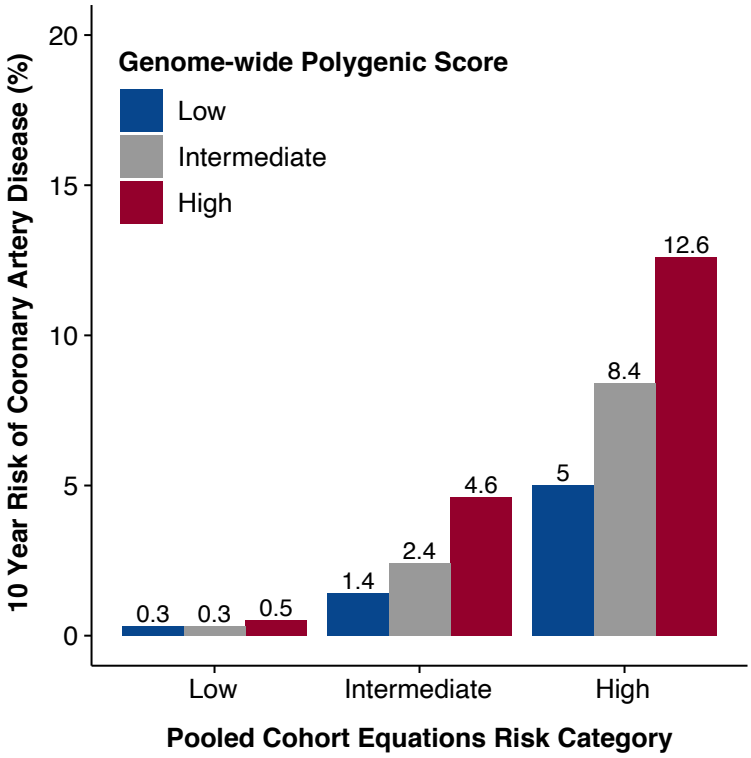
Scatter plot of the relationship between the genome-wide polygenic score for coronary artery disease and 10-year risk of atherosclerotic cardiovascular disease predicted by the Pooled Cohort Equations in 5,685 participants of the Malmö Diet and Cancer Cardiovascular Cohort. Pearson's coefficient was 0.03 and P=0.08.

Figure IV. 10-year and Lifetime Risk of Coronary Artery Disease According to Pooled Cohort Equations Risk Category and Genome-wide Polygenic Score in the Malmö Diet and Cancer Cardiovascular Cohort



Participants of the Malmö Diet and Cancer Cardiovascular Cohort were first stratified into low (bottom quintile), intermediate (quintiles 2-4), and high (top quintile) 10-year risk of atherosclerotic cardiovascular disease categories using the Pooled Cohorts Equation clinical risk estimator. Next individuals were stratified into low (bottom quintile), intermediate (quintiles 2-4), or high (top quintile) polygenic risk according to the genome−wide polygenic score for coronary artery disease (GPS_{CAD}). **A**) 10-year risk of incident coronary artery disease is displayed, based on a Cox proportional hazard model using follow up time as the time scale and standardized for age and sex using the population mean. **B**) Lifetime risk of coronary artery disease − defined as cumulative risk by 90 years of age − is displayed, based on a Cox proportional hazard model using age as the underlying time scale and standardized for sex using the population mean.

Figure V. 10-year Risk of Coronary Artery Disease According to Pooled Cohort Equations Risk Category and Genome-wide Polygenic Score in the UK Biobank



Participants of the UK Biobank were first stratified into low (bottom quintile), intermediate (quintiles 2-4), and high (top quintile) 10-year risk of atherosclerotic cardiovascular disease categories using the Pooled Cohorts Equation clinical risk estimator. Next individuals were stratified into low (bottom quintile), intermediate (quintiles 2-4), or high (top quintile) polygenic risk according to the genome-wide polygenic score for coronary artery disease (GPS_{CAD}). The 10-year risk of incident coronary artery disease is displayed, based on a Cox proportional hazard model using follow up time as the time scale and standardized for age and sex using the population mean.

Table I. Baseline Characteristics of the Malmö Diet and Cancer Study Participants, Stratified by the Genome-wide Polygenic Score

	Decile 1 (n= 2,856)	**Deciles 2-9 (n= 22,844)**	**Decile 10 (n= 2,856)**	P*
Age (years)	58.2 ± 7.7	57.9 ± 7.6	57.5 ± 7.4	6×10^{-4}
Male Sex, n (%)	1,167 (40.9)	8,841 (38.7)	1,055 (36.9)	9×10^{-3}
European ancestry, n (%)	2,820 (98.7)	22,642 (99.1)	2,824 (98.9)	0.09
Current Smoker, n (%)	724 (27.1)	6,134 (28.6)	1,552 (28.1)	0.25
Family History of CAD, n (%)	755 (26.4)	7,913 (34.6)	1,208 (42.3)	3×10^{-35}
Body Mass Index (kg/m^2)	25.6 ± 3.8	25.8 ± 4.1	25.9 ± 4.0	6×10^{-4}
Systolic Blood Pressure (mmHg)	140 ± 20	141 ± 20	142 ± 20	6×10^{-5}
Diastolic Blood Pressure (mmHg)	85 ± 10	86 ± 10	86 ± 10	7×10^{-3}
Use of anti-hypertensives, n (%)	416 (14.6)	3,516 (15.4)	982 (17.6)	3×10^{-3}
Use of lipid-lowering, n (%)	38 (1.3)	465 (2.0)	109 (3.8)	3×10^{-11}
Diabetes, n (%)	121 (4.2)	987 (4.3)	128 (4.5)	0.89
Apolipoprotein B (mg/dL)	102.0 ± 25.6	107.2 ± 26.0	111.7 ± 26.6	1×10^{-41}
Apolipoprotein A1 (mg/dL)	158.1 ± 27.7	157.1 ± 28.3	156.3 ± 28.1	2×10^{-2}

Individuals were stratified into low (decile 1), intermediate (deciles 2-9), or high (decile 10) polygenic risk according to the genome-wide polygenic score for coronary artery disease (GPS_{CAD})

*P-values were obtained from ANOVA for continuous variables and Chi squared test for categorical variables

Table II. Baseline Characteristics of the Malmö Diet and Cancer Cardiovascular Cohort (n=5,685) by Incident Coronary Artery Disease Status.

	Incident coronary artery disease (n= 815)	No incident coronary artery disease (n= 4,870)	P*
Age (years)	59.1 ± 5.6	57.1± 5.9	3×10^{-31}
Male Sex, n (%)	477 (58.5)	1,863 (38.3)	8×10^{-34}
European ancestry, n (%)	810 (99.4)	4,830 (99.2)	0.60
Current Smoker, n (%)	262 (32.1)	1,256 (25.8)	3×10^{-11}
Family History of CAD, n (%)	348 (42.7)	1,709 (35.1)	2×10^{-06}
Body Mass Index (kg/m^2)	26.6 ± 4.2	25.6± 3.9	1×10^{-09}
Systolic Blood Pressure (mmHg)	147.9 ± 19.2	140.1 ± 18.8	8×10^{-19}
Diastolic Blood Pressure (mmHg)	89.6 ± 9.5	86.4 ± 9.4	6×10^{-12}
Use of anti−hypertensives, n (%)	175 (21.5)	668 (13.7)	3×10^{-08}
Use of lipid−lowering, n (%)	27 (3.3)	72 (1.4)	5×10^{-04}
Diabetes, n (%)	80 (9.8)	177 (3.6)	2×10^{-13}
Total cholesterol (mg/dL)	245.0 ± 42.0	237.6 ± 42.4	1×10^{-05}
LDL cholesterol (mg/dL)	168.5 ± 37.6	159.9 ± 38.1	2×10^{-06}
HDL cholesterol (mg/dL)	49.0 ± 13.6	54.3 ± 14.4	2×10^{-10}
Triglycerides (mg/dL)	138.7 ± 84.3	118.2 ± 67.9	5×10^{-09}
Apolipoprotein B (mg/dL)	119.0 ± 26.9	109.9 ± 26.3	1×10^{-13}
Apolipoprotein A1 (mg/dL)	153.2 ± 28.7	161.6 ± 29.4	2×10^{-8}

*P-values were obtained from Cox proportional hazard models adjusted for age and sex

Table III. Multivariable Cox Regression Model including Genome-wide Polygenic Score for Prediction of Incident Coronary Artery Disease Events in the Malmö Diet and Cancer Cardiovascular Cohort

	HR	95% CI	P
Age (10-Year)	1.91	1.66−1.20	3×10^{-19}
Sex (Male)	2.18	1.85−2.58	5×10^{-20}
Family history of CAD	1.37	1.18−1.60	5×10^{-5}
Current Smoker	1.80	1.53−2.12	3×10^{-12}
Diabetes	2.16	1.65−2.83	2×10^{-8}
Body Mass Index (SD)	1.08	0.99−1.17	8×10^{-2}
Systolic Blood Pressure (SD)	1.29	1.19−1.40	4×10^{-10}
HDL cholesterol (SD)	0.82	0.75−0.90	4×10^{-5}
Total Cholesterol (SD)	1.12	1.03−1.21	7×10^{-3}
GPS_{CAD} (SD)	1.45	1.34−1.56	2×10^{-22}

The Malmö Diet and Cancer Cardiovascular Cohort included 4,937 individuals with 688 cases of coronary artery disease (CAD) reported during follow-up.
HR, hazard ratio; CI, confidence interval; SD, standard deviation.

Table IV. Correlation between the Genome-wide Polygenic Score and Individual Clinical Risk Factors in the Malmö Diet and Cancer Study

Risk Factor	Pearson's r	P-value
Body Mass Index	0.01	0.44
Systolic Blood Pressure	0.04	0.002
Apolipoprotein B	0.10	1×10^{-13}
Apolipoprotein A1	-0.02	0.24
Total Cholesterol	0.08	1×10^{-8}
LDL Cholesterol	0.10	1×10^{-12}
HDL Cholesterol	-0.02	0.17

The correlations between the genome-wide polygenic score adjusted for top 10 principal components and traditional risk factors were assessed using Pearson's correlation in the Malmö Diet and Cancer Cardiovascular Cohort (n = 5,685).

Table V. 10-year and Lifetime Risk of Coronary Artery Disease According to Pooled Cohort Equations Risk Category and Genome-wide Polygenic Score in the Malmö Diet and Cancer Cardiovascular Cohort

			10-Year Risk*		Lifetime Risk†	
10-yr ASCVD Risk by PCE	**GPS_{CAD}**	**N_{Total}**	**N_{CAD}**	**Risk (95% CI)**	**N_{CAD}**	**Risk (95% CI)**
Low (< 5%)	Low	348	2	1.1 (0.5−1.6)	15	8.7 (4.3−12.8)
	Intermediate	1,041	6	1.3 (0.9−1.7)	54	10.4 (7.6−13.0)
	High	309	5	2.2 (1.4−3.1)	40	17.5 (11.1−23.3)
Borderline (>= 5% & < 7.5%)	Low	172	1	1.7 (0.7−2.8)	11	11.3 (4.8−17.4)
	Intermediate	485	3	1.8 (1.2−2.5)	33	11.6 (7.8−15.3)
	High	161	5	5.9 (3.8−7.9)	34	34.1 (24.1−42.8)
Intermediate (>= 7.5% & < 20%)	Low	365	12	3.3 (2.2−4.4)	40	18.1 (12.8−23.1)
	Intermediate	1,098	56	5.4 (4.4−6.3)	192	28.4 (24.7−31.9)
	High	388	38	8.4 (6.5−10.2)	100	41.3 (34.6−47.3)
High (>= 20%)	Low	110	8	7.2 (4.0−10.2)	20	32.4 (19.6−43.1)
	Intermediate	341	50	12.4 (9.8−14.8)	107	50.4 (43.1−56.8)
	High	123	27	19.2 (14.2−23.9)	56	67.3 (55.9−75.8)

*10-year risk of incident coronary artery disease is displayed, based on a Cox proportional hazard model using follow up time as the time scale and standardized for age and sex using the population mean.
†Lifetime risk of coronary artery disease − defined as cumulative risk by 90 years of age − is displayed, based on a Cox proportional hazard models using age as the underlying time scale and standardized sex using the population mean.
ASCVD, Atherosclerotic cardiovascular disease; GPS_{CAD}; PCE, Pooled Cohort Equations Genome−wide Polygenic Score; CAD, Coronary Artery Disease; CI, Confidence Interval

Table VI. Net Reclassification Improvement after Addition of the Genome-wide Polygenic Score to the Pooled Cohort Equations in the Malmö Diet and Cancer Study

		PCE + GPS_{CAD}			
	Events	< 5%	>= 5% and < 7.5%	>= 7.5% and < 20%	>= 20%
PCE	< 5%	54	18	10	0
	>= 5% and < 7.5%	5	13	18	0
	>= 7.5% and < 20%	3	9	58	9
	>= 20%	0	0	1	15
		PCE + GPS_{CAD}			
	Non-Events	< 5%	>= 5% and < 7.5%	>= 7.5% and < 20%	>= 20%
PCE	< 5%	3197	208	33	0
	>= 5% and < 7.5%	188	184	124	0
	>= 7.5% and < 20%	26	111	306	25
	>= 20%	0	1	12	28
Net Reclassification Improvement (NRI)					
NRI Events (95% CI)		0.173 (0.088 to 0.199)			
NRI Non-Events (95% CI)		-0.009 (-0.018 to -0.002)			
NRI		0.165 (0.076 to 0.182)			

The net reclassification improvement (NRI) was obtained from predicted 10-year risks from Cox proportional hazard models.
PCE, Pooled Cohort Equations; GPS_{CAD}, Genome-wide Polygenic Score; CAD, Coronary Artery Disease; CI, Confidence Interval

Table VII. Baseline Characteristics of the UK Biobank (n=325,003) by Incident Coronary Artery Disease Status

	Incident coronary artery disease (n= 7,708)	No incident coronary artery disease (n= 317,295)	P*
Age (years)	61.1 ± 6.6	56.7 ± 9.1	0
Male Sex, n (%)	5,535 (71.8)	138,003 (43.5)	0
Self-reported Ancestry			
European, n (%)	7,189 (93.3)	297,081 (93.6)	1×10^{-15}
South Asian, n (%)	277 (3.6)	6,555 (2.1)	1×10^{-35}
African, n (%)	73 (0.9)	5,687 (1.8)	0.05
Chinese, n (%)	13 (0.2)	1104 (0.3)	0.24
Other, n (%)	156 (2.0)	6868 (2.2)	0.01
Current Smoker, n (%)	262 (15.7)	29,487 (9.3)	1×10^{-87}
Family History of CAD, n (%)	348 (52.9)	132,783 (41.8)	6×10^{-80}
Body Mass Index (kg/m^2)	28.5 ± 4.6	27.3 ± 4.7	1×10^{-90}
Systolic Blood Pressure (mmHg)	148.0 ± 20.2	139.4 ± 19.6	3×10^{-67}
Diastolic Blood Pressure (mmHg)	84.7 ± 11.2	82.2 ± 10.7	1×10^{-24}
Use of anti−hypertensives, n (%)	1,198 (15.5)	32,333 (10.2)	1×10^{-17}
Use of lipid−lowering, n (%)	2,398 (31.1)	45,896 (14.5)	3×10^{-93}
Prevalent Diabetes, n (%)	977 (12.7)	14,825 (4.7)	4×10^{-116}
Total cholesterol (mg/dL)	221.8 ± 49.8	221.7 ± 43.3	9×10^{-16}
LDL cholesterol (mg/dL)	141.8 ± 38.1	138.5 ± 33.0	1×10^{-36}
HDL cholesterol (mg/dL)	49.7 ± 12.9	56.6 ± 14.8	2×10^{-131}
Triglycerides (mg/dL)	188.8 ± 107.3	152.6 ± 89.5	6×10^{-120}
Apolipoprotein B (mg/dL)	108.4 ± 26.5	103.6 ± 23.5	5×10^{-71}
Apolipoprotein A1 (mg/dL)	143.72 ± 24.8	154.7 ± 27.0	2×10^{-110}
GPS_{CAD} (SD)	0.39 ± 1.0	−0.01 ± 1	5×10^{-303}

*P-values were obtained from Cox proportional hazard models adjusted for age and sex

Table VIII. Multivariable Cox Regression Model Including Genome-wide Polygenic Score for Prediction of Incident Coronary Artery Disease Events in the UK Biobank

	HR	95% CI	P
Age (10-Year)	2.14	2.05−2.22	8×10^{-307}
Sex (Male)	2.66	2.50−2.83	1×10^{-205}
Family history of heart disease	1.41	1.34−1.48	1×10^{-38}
Current Smoker	1.79	1.67−1.92	1×10^{-56}
Diabetes	2.05	1.88−2.23	7×10^{-62}
Body Mass Index (SD)	1.07	1.04−1.10	5×10^{-6}
Systolic Blood Pressure (SD)	1.19	1.16−1.23	9×10^{-40}
HDL cholesterol (SD)	0.70	0.67−0.72	2×10^{-83}
Total Cholesterol (SD)	1.25	1.22−1.28	7×10^{-59}
GPS_{CAD} (SD)	1.46	1.42−1.50	5×10^{-184}

The UK Biobank included 260,282 individuals with 5,935 cases of coronary artery disease (CAD) reported during follow-up.
HR, hazard ratio; CI, confidence interval; SD, standard deviation.

Table IX. Risk of Coronary Artery Disease According to Pooled Cohort Equations Risk Category and Genome-wide Polygenic Score in the UK Biobank

10-yr ASCVD by PCE	GPS_{CAD}	N_{Total}	N_{CAD}	10-Year Risk* (95% CI)
Low (< 5%)	Low	25,015	86	0.5 (0.4−0.6)
	Intermediate	73,152	348	0.7 (0.6−0.8)
	High	24,037	241	1.4 (1.2−1.6)
Borderline (>= 5% & < 7.5%)	Low	7,604	64	1.2 (0.9−1.5)
	Intermediate	23,357	372	2.3 (1.9−1.5)
	High	8,045	250	4.4 (3.7−5.1)
Intermediate (>= 7.5% & < 20%)	Low	18,298	358	2.8 (2.4−3.2)
	Intermediate	55,712	1,877	4.8 (4.3−5.4)
	High	18,680	1,088	8.3 (7.3−9.3)
High (>= 20%)	Low	5,245	225	6.3 (5.2−7.3)
	Intermediate	16,603	1,196	10.4 (9.2−11.6)
	High	5,439	571	15.1 (13.2−16.9)

*Risk of incident coronary artery disease is displayed, based on a Cox proportional hazard model using follow up time as the time scale and standardized for age and sex using the population mean.
ASCVD, Atherosclerotic cardiovascular disease; PCE, Pooled Cohort Equations; GPS_{CAD}, Genome-wide Polygenic Score; CAD, Coronary Artery Disease; CI, Confidence Interval

Table X. Net Reclassification Improvement after Addition of the Genome-wide Polygenic Score to the Pooled Cohort Equations in the UK Biobank

		PCE + GPS_{CAD}			
	Events	< 5%	>= 5% and < 7.5%	>= 7.5% and < 20%	>= 20%
PCE	< 5%	3,573	463	144	0
	>= 5% and < 7.5%	296	451	380	2
	>= 7.5% and < 20%	34	145	612	61
	>= 20%	0	1	8	50
		PCE + GPS_{CAD}			
	Non-Events	< 5%	>= 5% and < 7.5%	>= 7.5% and < 20%	>= 20%
PCE	< 5%	133,960	3,679	729	0
	>= 5% and < 7.5%	3,771	3,499	2,078	5
	>= 7.5% and < 20%	443	1,350	3,081	181
	>= 20%	0	0	60	91
Net Reclassification Improvement (NRI)					
NRI Events (95% CI)		0.091 (0.077 to 0.105)			
NRI Non-Events (95% CI)		-0.006 (-0.007 to -0.006)			
NRI		0.085 (0.071 to 0.098)			

The net reclassification improvement (NRI) was obtained from predicted 8-year risks from Cox proportional hazard models.
PCE, Pooled Cohort Equations; GPS_{CAD}, Genome-wide Polygenic Score; CAD, Coronary Artery Disease; CI, Confidence Interval

Table XI. Association of the Genome-wide Polygenic Score with Clinical Risk Factors in the UK Biobank

	Low (n= 65,001)	Intermediate (n= 195,001)	High (n= 65,001)	P*
Age (years)	57.0 ± 8.1	56.9 ± 8.1	56.6 ± 8.1	4×10^{-21}
Male Sex, n (%)	29,232 (45.0)	86,296 (44.3)	28,010 (43.1)	1×10^{-13}
Self-reported Ancestry				
European, n (%)	60,873 (93.6)	182,764 (93.7)	60,633 (93.3)	0.77
South Asian, n (%)	1,452 (2.2)	3,922 (2.0)	1,458 (2.2)	0.23
African, n (%)	1,069 (1.6)	3,574 (1.8)	1,117 (1.7)	0.57
Chinese, n (%)	247 (0.4)	673 (0.3)	197 (0.3)	9×10^{-5}
Other, n (%)	1,360 (2.1)	4,068 (2.1)	1,596 (2.5)	0.005
Current Smoker, n (%)	5,989 (9.2)	18,298 (9.4)	6,411 (9.9)	2×10^{-5}
Family History of CAD, n (%)	23,992 (36.9)	81,707 (41.9)	31,164 (47.9)	$<1\times10^{-100}$
Body Mass Index (kg/m^2)	27.1 ± 4.7	27.3 ± 4.7	27.5 ± 4.8	1×10^{-52}
Systolic Blood Pressure (mmHg)	139 ± 20	140 ± 20	141 ± 20	$<1\times10^{-100}$
Diastolic Blood Pressure (mmHg)	82 ± 11	82 ± 11	83 ± 11	4×10^{-87}
Use of anti−hypertensives, n (%)	6,460 (9.9)	20,236 (10.4)	6,835 (10.5)	4×10^{-6}
Use of lipid−lowering, n (%)	7,774 (12.0)	28,844 (14.8)	11,676 (18.0)	$<1\times10^{-100}$
Prevalent Diabetes, n (%)	2,832 (4.4)	9,471 (4.9)	3,499 (5.5)	2×10^{-26}
Total cholesterol (mg/dL)	218.3 ± 42.2	221.8 ± 43.4	225.0 ± 44.7	$<1\times10^{-100}$
LDL cholesterol (mg/dL)	135.4 ± 32.2	138.7 ± 33.1	141.7 ± 34.1	$<1\times10^{-100}$
HDL cholesterol (mg/dL)	57.1 ± 15.0	56.4 ± 14.8	56.0 ± 14.6	1×10^{-60}
Apolipoprotein B (mg/dL)	100.9 ± 23.1	103.8 ± 23.6	106.4 ± 24.0	$<1\times10^{-100}$
Apolipoprotein A1 (mg/dL)	155.5 ± 27.3	154.4 ± 26.9	153.6 ± 27.0	2×10^{-51}
Triglycerides (mg/dL)	128.0 (90.3-186.6)	130.6 (92.1-189.3)	131.2 (92.6-189.9)	4×10^{-34}
C-Reactive Protein (mg/L)	1.28 (0.63-2.64)	1.31 (0.65-2.71)	1.32 (0.65-2.57)	3×10^{-14}
Preeclampsia (Females), n (%)	315 (0.97)	1,222 (1.12)	411 (1.13)	0.001
Chronic Kidney Disease, n (%)	1,071 (1.6)	3,392 (1.7)	1,136 (1.7)	0.001
Rheumatoid Arthritis, n (%)	1,106 (1.7)	3,623 (1.9)	1,220 (1.9)	0.006
Psoriasis, n (%)	1,004 (1.5)	3,074 (1.6)	1,042 (1.6)	0.15
Inflammatory Bowel Disease, n (%)	906 (1.4)	2684 (1.4)	956 (1.5)	0.18

Individuals were stratified into low (quintile 1), intermediate (quintiles 2-4), or high (quintile 5) polygenic risk according to the genome-wide polygenic score for coronary artery disease (GPS_{CAD}). Data is presented as mean ± standard deviation for normally distributed continuous variables and as median (Interquartile range) for skewed continuous variables.
*P-values were obtained from regression models adjusted for age and sex. Skewed continuous variables were natural log-transformed.

Table XII. Discriminative Capacity of the Genome-wide Polygenic Score, the Pooled Cohorts Equation 10-year Risk Estimate, and Both Within the Atherosclerosis Risk in Communities Study

Predictors	C-statistic	95%CI
Age, sex, first four principal components of ancestry	0.672	0.656−0.688
+ GPS_{CAD}	0.697	0.681−0.713
+ PCE 10-year risk estimate	0.726	0.710−0.741
+ GPS_{CAD} and PCE 10-year risk estimate	0.739	0.723−0.755

Among 7,318 white participants of the Atherosclerosis Risk in Communities (ARIC) study free of coronary artery disease (CAD) at time of enrollment, 1,119 incident coronary artery disease events were observed over a median follow-up of 18.9 (IQR 17.3 to 19.5 years). The CAD endpoint was a composite inclusive of myocardial infarction, coronary revascularization, and death from coronary causes. Harrell's C-statistic was computed within Cox proportional hazard models for each set of predictors.
GPS, genome-wide polygenic score; PCE, Pooled Cohort Equations; CI, confidence interval

Regeneron Genetics Center Banner Author List and Contribution Statements

All authors/contributors are listed in alphabetical order.

RGC Management and Leadership Team
Goncalo Abecasis, Ph.D., Aris Baras, M.D., Michael Cantor, M.D., Giovanni Coppola, M.D., Aris Economides, Ph.D., Luca A. Lotta, M.D., Ph.D., John D. Overton, Ph.D., Jeffrey G. Reid, Ph.D., Alan Shuldiner, M.D.

Contribution: All authors contributed to securing funding, study design and oversight. All authors reviewed the final version of the manuscript.

Sequencing and Lab Operations
Christina Beechert, Caitlin Forsythe, M.S., Erin D. Fuller, Zhenhua Gu, M.S., Michael Lattari, Alexander Lopez, M.S., John D. Overton, Ph.D., Thomas D. Schleicher, M.S., Maria Sotiropoulos Padilla, M.S., Karina Toledo, Louis Widom, Sarah E. Wolf, M.S., Manasi Pradhan, M.S., Kia Manoochehri, Ricardo H. Ulloa.

Contribution: C.B., C.F., K.T., A.L., and J.D.O. performed and are responsible for sample genotyping. C.B, C.F., E.D.F., M.L., M.S.P., K.T., L.W., S.E.W., A.L., and J.D.O. performed and are responsible for exome sequencing. T.D.S., Z.G., A.L., and J.D.O. conceived and are responsible for laboratory automation. M.P., K.M., R.U., and J.D.O are responsible for sample tracking and the library information management system.

Genome Informatics
Xiaodong Bai, Ph.D., Suganthi Balasubramanian, Ph.D., Leland Barnard, Ph.D., Andrew Blumenfeld, Gisu Eom, Lukas Habegger, Ph.D., Alicia Hawes, B.S., Shareef Khalid, Jeffrey G. Reid, Ph.D., Evan K. Maxwell, Ph.D., William Salerno, Ph.D., Jeffrey C. Staples, Ph.D., Ashish Yadav, M.S.

Contribution: X.B., A.H., W.S. and J.G.R. performed and are responsible for analysis needed to produce exome and genotype data. G.E. and J.G.R. provided compute infrastructure development and operational support. S.K., S.B., and J.G.R. provide variant and gene annotations and their functional interpretation of variants. E.M., L.B., J.S., A.B., A.Y., L.H., J.G.R. conceived and are responsible for creating, developing, and deploying analysis platforms and computational methods for analyzing genomic data.

Research Program Management
Marcus B. Jones, Ph.D., Lyndon J. Mitnaul, Ph.D.

Contribution: All authors contributed to the management and coordination of all research activities, planning and execution. All authors contributed to the review process for the final version of the manuscript.

Major Resources Table

In order to allow validation and replication of experiments, all essential research materials listed in the Methods should be included in the Major Resources Table below. Authors are encouraged to use public repositories for protocols, data, code, and other materials and provide persistent identifiers and/or links to repositories when available. Authors may add or delete rows as needed.

Animals (in vivo studies)

Species	Vendor or Source	Background Strain	Sex	Persistent ID / URL
NA	NA	NA	NA	NA

Genetically Modified Animals

	Species	Vendor or Source	Background Strain	Other Information	Persistent ID / URL
Parent - Male	NA	NA	NA	NA	NA
Parent - Female	NA	NA	NA	NA	NA

Antibodies

Target antigen	Vendor or Source	Catalog #	Working concentration	Lot # (preferred but not required)	Persistent ID / URL
NA	NA	NA	NA	NA	NA

DNA/cDNA Clones

Clone Name	Sequence	Source / Repository	Persistent ID / URL
NA	NA	NA	NA

Cultured Cells

Name	Vendor or Source	Sex (F, M, or unknown)	Persistent ID / URL
NA	NA	NA	NA

Data & Code Availability

Description	Source / Repository	Persistent ID / URL
NA	NA	NA

Other

Description	Source / Repository	Persistent ID / URL
NA	NA	NA